

The Grid Analysis Environment (GAE): A Vision of Distributed, Collaborative Environments for Scientific Data Analysis.

Dimitri Bourilkov^{*}, Julian Bunn[♦], Rick Cavanaugh^{*}, Iosif Legrand^{*}, Harvey Newman^{*}, Conrad Steenberg^{*}, Michael Thomas^{*}, Frank van Lingen[♦]

As grid middleware matures and grid (web) services become more prolific the development of "higher level" services that can take intelligent decisions based on these "lower level" services is required. In 2003, Caltech and the University of Florida initiated the Grid Analysis Environment (GAE) project with the goals of developing, integrating, and deploying a web service framework and services for physics analysis (especially CMS[1]) at the LHC. The project's kernel is a Grid-based portal called Clarens [4]. GAE is a continuation of the CAIGEE project [7].

The GAE will be used by a large, diverse community. It will need to support hundreds, even thousands, of analysis tasks with widely varying requirements. It will need to employ priority schemes, and robust authentication and security mechanisms. Most challenging, the GAE will need to operate well in what we expect to be a severely resource limited global computing system. The GAE is where the critical physics analysis gets done, where the Grid end-to-end services are exposed to a very demanding physicist *clientele*, who will have to learn how to collaborate at large distances on challenging analysis topics.

The LHC experiments' computing models were initially based on a hierarchical organization of Tier 0, multiple Tier 1 and multiple Tier 2, centres. Most of the data will "flow" from the Tier 0, to the Tier 1s and then on to the Tier 2s and downstream. The Tier 0 and Tier 1 centres are powerful in terms of CPU power, data storage, and bandwidth. This hierarchical model is such that the sum of resources below a certain Tier-N are of approximately the same amount as that located at the Tier-N [2]. Data is organized such that the institutes (represented by tiers) will be (in most cases) physically close to the data they want to analyze.

The hierarchical model is the basis on which data will be distributed. However, the unpredictable patterns of physics analysis as a whole will lead to data and jobs being moved around between different peers in an essentially chaotic fashion. Although it is possible to make predictions on how best to organize data, the unpredictable, chaotic analysis patterns tend to defeat simple rationalization. Depending on how "hot" (popular) or "cold" (infrequently requested) data is, attention will shift between the different data sets. Furthermore, multiple geographically distant users might be interested in data that is geographically dispersed among the different tiers. Because of this, data will need to be replicated. And so it will not always be possible to move data around so that it fits into a strict hierarchical model.

These data movements and job movements outside the hierarchical model, although relative small compared to the data movement within the hierarchical model, will be significant consumers of resources (network, CPU, storage). When users submit jobs to a peer, middleware will discover which resources are available on other peers that can satisfy the job requests. Although a large number of job requests will fit with the hierarchical model, others will not. The more powerful peers (super peers) will receive more job and data requests, and will host a wider variety of services.

Additionally, monitoring the GAE as it is used is of crucial importance in order to make intelligent decisions on data distribution, job execution and data management. The MonaLISA [3] monitoring

^{*} University of Florida

[♦] California Institute of Technology

framework is capable of satisfying this need, and it has been deployed on many sites to monitor disk, and CPU usage. In the future it will be able to perform complex trend analysis on GAE resource access patterns, and so eventually enable automatic resource co-scheduling.

Data in the GAE will be distributed among multiple Tier-N centers: it is important that physicists can gain access quickly and reliably to that data. The development of robust and stable "ultra scale" (100 Gbits/sec and higher) networking in the wide area, is therefore critical. This new generation of ultra scale computing with Petabyte- and in some cases Exabyte- sized datasets promises to drive discoveries in fundamental and applied sciences of the next decade [5].

Within most scientific disciplines, collaborations are vital in making new discoveries: ease of human interactions with experts and peers in the field is essential. Since much of the time these experts and peers are not all located at the same place, the need for collaborative systems is paramount. The current VRVS system [6] allows geographically dispersed groups of scientists to schedule and meet in, virtual rooms and interact in a way that is difficult or impossible otherwise.

In Summary, we have described a highly dynamic and complex environment in which it is crucial that "high level" (=intelligent) services take as many decisions on what to execute where and by whom that they can. These decisions will be based on (trend) analysis of monitoring data from a great variety of resources. Services such as schedulers and replica managers will implement the decisions using the available information so as to ensure efficient and "fair" use of Grid resources [8] for the large user community. Ideally, human intervention should be minimized to prevent the '1000 system administrators' support nightmare that will otherwise prevail. But intelligent services alone are not sufficient: fast and reliable data transfer is vital to enable physicists to perform analysis "at the touch of a button". Although Grid services are becoming more complex and sophisticated there will always be the need for one or more humans to remain in control!

References:

- [1] CMS: <http://cmsdoc.cern.ch/cms.html>
- [2] J. Bunn, H. Newman, "*Data Intensive Grids for High Energy Physics*", p859-p906, in "Grid Computing", edited by F. Berman, G. Fox, A. Hey, 2003, Wiley
- [3] Legrand, I., Newman, H., Galvez, P., Voicu, E., Cirstoiu, C., "*MonaLISA: A Distributed Monitoring Service Architecture*", Computing for High Energy Physics, La Jolla, California, 2003
- [4] Steenberg, C., Aslakson, E., Bunn, J., Newman, H., Thomas, M., Van Lingen, F., "*The Clarens Web Service Architecture*", Computing for High Energy Physics, La Jolla, California, 2003
- [5] C. Jin, D. Wei, S. H. Low, G. Buhrmaster, J. Bunn, D. H. Choe, R. L. A. Cottrell, J. C. Doyle, W. Feng, O. Martin, H. Newman, F. Paganini, S. Ravot, S. Singh, "*FAST TCP: From Theory to Experiments*" submitted for publication, April 1, 2003
- [6] VRVS: <http://www.vrvs.org/>
- [7] Newman, H. Branson, J. "*CMS Analysis: an Interactive Grid-Enabled Environment (CAIGEE)*" Submitted to the 2002 NSF Information and Technology Research Program Proposal #6116240
- [8] In, J, Avery, P., Cavanaugh, R., Ranka, S. "*Policy Based Scheduling for Simple Quality of Service in a Grid*," Proceedings of the International Parallel & Distributed Processing Symposium (IPDPS), Santa Fe, New Mexico, 2004